DOCUMENT RESUME

ED 289 505 IR 012 972

AUTHOR Ellington, Henry
TITLE Short Answer Questions. Teaching and Learning in Higher Education, 22.
INSTITUTION Scottish Central Institutions Committee for Educational Development.
SPONS AGENCY Robert Gordon's Inst. of Technology, Aberdeen (Scotland).
PUB DATE 87
NOTE 14p.; For a related guide, see IR 012 970.
PUB TYPE Guides - Non-Classroom Use (055) -- Information Analyses (070)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Foreign Countries; Higher Education; Instructional Material Evaluation; *Questioning Techniques; Scoring; *Test Construction; *Test Format; *Test Items
IDENTIFIERS *Short Answer Tests

ABSTRACT

The second of three sequels to the booklet "Student Assessment," this booklet begins by describing and giving examples of three different forms that short-answer questions can take: (1) completion items; (2) unique-answer questions; and (3) open short-answer questions. Guidelines are then provided for deciding which type of question to use in a given situation, and the task of writing short-answer questions is broken down into three stages and described in detail. Methods for the evaluation of short-answer questions are also discussed, including evaluation by a colleague or validation panel and quantitative evaluation. Advice on how to mark short-answer questions concludes the booklet. Two general references and three subject-based references for teachers of chemistry, physics, and mathematics are listed. (MES)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Henry Ellington

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

This booklet was first published internally in Robert Gordon's Institute of Technology, Aberdeen as part of the Institute's staff development programme. It was written by Dr Henry Ellington of RGIT's Educational Technology Unit.

CICED gratefully acknowledges the co-operation of RGIT and the author in the publication of the present edition of the booklet.

© CICED 1987

Short-Answer Questions

Introduction

This booklet forms a sequel to the booklet on 'Student assessment', which presents a broad survey of the field and examines the different assessment techniques that can be employed therein. The present booklet deals in detail with one particular group of assessment instruments, namely, *short answer questions*.

The booklet begins by examining the different forms that short-answer questions can take and offering guidance on how to decide which type to use in a given situation. Next, it gives detailed guidance on how to write short-answer questions of different types. Finally, it offers practical advice on how to evaluate and mark short-answer questions.

The different types of short-answer questions

As was described in the booklet on 'Student assessment', short-answer questions are similar to multiple-choice questions in that a clearly-defined answer is required, but differ from the latter in that the answer has to be supplied by the testee rather than chosen from a number of options provided. Such questions are often classed as *objective questions*, because they can (if suitably designed) be marked with a very high degree of reliability. Strictly speaking, however, they are not true objective questions in the same way that the various forms of multiple-choice questions are, since the marker may sometimes have to exercise a certain amount of subjective judgment in deciding whether an answer is satisfactory. The marker may, for example, have to decide whether a wrongly-spelt word, a partly correct answer, or a perfectly good answer that the question setter did not anticipate is acceptable, and may also have to decide whether an explanation or description is satisfactory. Obviously, the amount of subjective judgment involved tends to increase with the length and complexity of the required answer, while the degree of 'objectivity' shows a corresponding fall. Let us now take a look at some of the most common forms that short-answer questions can take.

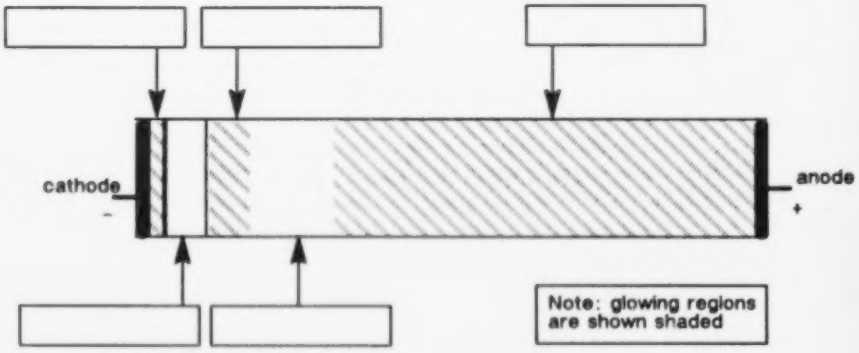
Completion items

In their simplest form, these consist of incomplete statements, the testee having to supply the missing words, terms, symbols, etc. Two typical examples are shown below.

Example 1 (a simple completion item where only a single answer has to be supplied)
'When a small piece of sodium metal is dropped into water, [ ] is given off in gaseous form'. (hydrogen)

Example 2 (a more complex completion item where several answers have to be supplied)
'According to the historian and philosopher of science Thomas Kuhn, a branch or sub-branch of science only gains maturity once it acquires a [ ]. Once it does so, it generally enters a period of stability and steady, albeit unspectacular, progress known as [ ]. The evolution of a mature branch or sub-branch of science generally consists of a succession of such stable pericds, interspersed with shorter periods of rapid change known as [ ].'
(paradigm; normal science; extraordinary science OR scientific revolutions).
Completion items can also be built round things like maps, diagrams, drawings and photographs, with the testee again having to supply missing pieces of information. Two examples of this genre are given below.

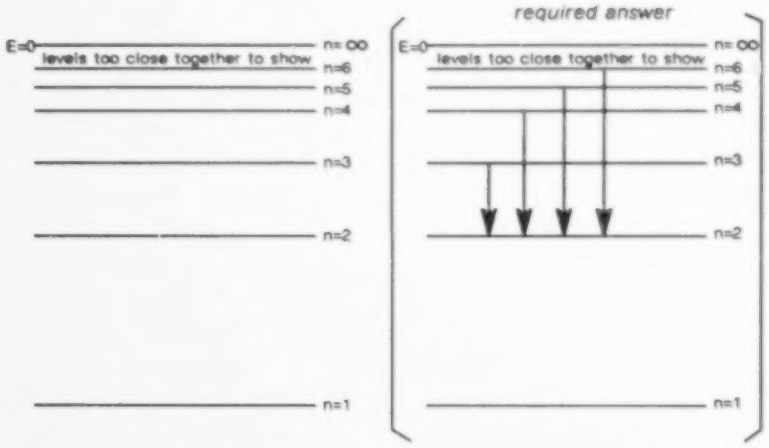
Example 3 (a completion item that involves adding verbal information to graphical material).
'The following diagram shows the five main regions of a *glow discharge*. Write the names of these different regions in the spaces provided.'



(Answers (left to right): cathode glow, cathode dark space, negative glow, Faraday dark space, positive column).

Example 4 (a completion item that involves adding graphical information to graphical material).

'The electronic energy level diagram for atomic hydrogen is shown below. Draw arrows on this diagram to represent the transitions that correspond to the first four emission lines of the Balmer series.'



Unique-answer questions

These take the form of actual questions (or instructions that imply questions), with the testee having to supply the answer. Such questions can themselves take a wide range of forms, some of the possibilities being shown below.

Example 1 (a simple unique-answer question)

'State the chemical formula of calcium bromide' ($CaBr_2$)

Example 2 (a unique answer question involving the interpretation of data and the performance of simple calculations based on same)

'A sample of radioactive material is found to have an activity of 256 microcuries at 1.00 PM, 64 microcuries at 2.00 PM and 1 microcurie at 5.00 PM

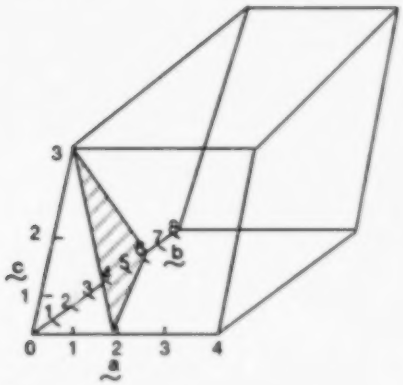
(a) What is the half life of the material?

(b) What is its activity at 4.00 PM?'

(answers: 30 min.; 4 microcuries)

6

Example 3 (a unique answer question involving interpretation of graphical information)



'The above figure represents a crystal unit cell having axial lengths a, b and c equal to 4 $\overset{\circ}{A}$, 8 $\overset{\circ}{A}$ and 3$\overset{\circ}{A}$ respectively. State the Miller indices of the shaded plane in preferred (i.e. integer) form, showing the various stages of the process by which you determine them.'

(Model answer:	Axial lengths	:	4 $\overset{\circ}{A}$		8 $\overset{\circ}{A}$		3 $\overset{\circ}{A}$	SUCCESSIVE STAGES BY WHICH ANSWER IS ARRIVED AT
	Intercept lengths	:	2 $\overset{\circ}{A}$		6 $\overset{\circ}{A}$		3 $\overset{\circ}{A}$	
	Fractional intercepts	:	$\frac{1}{2}$		$\frac{3}{4}$		1	
	Miller indices	:	2		$\frac{4}{3}$		1	
	Preferred form of Miller indices	:	6		4		3	REQUIRED ANSWER)

'Open' short-answer questions

These are similar to unique-answer questions except that they allow for some variation in the nature of the answer, either in terms of its intrinsic content or in terms of the way in which it is expressed. Two typical examples are given below.

Example 1 (a question that has several acceptable answers)

'Name two organs of the body that are particularly susceptible to damage by ionising radiation'.

(possible answers include the gonads, the blood-producing bone marrow, the lymph glands, the small intestines and the retina of the eye).

Example 2 (a question in which the answer can be expressed in a number of ways)

'State the revolutionary assumption made by Planck in deriving his theoretical formula for the shape of the black body emission spectrum at a given temperature'. (Here, the testee is required to explain how Planck postulated that radiation is emitted in discrete packets – or *quanta* – rather than in a continuous manner, as had hitherto been supposed.)

As in the case of multiple-choice questions, a large number of variations on these basic question types are possible.

Deciding which type of question to use in a given situation

Like multiple-choice questions, short-answer questions are best suited for assessing learning outcomes in the lower-to-middle part of the *cognitive domain*, i.e. for testing *knowledge, comprehension, application and analysis*. They are not particularly well suited for testing *synthesis* and *evaluation* (the highest levels of the cognitive domain), although it is possible to write items in these areas given sufficient ingenuity. Again like multiple-choice questions, they are of little use in testing non-cognitive skills such as communication skills, interpersonal skills and psychomotor skills. Thus, the first thing that anyone thinking of making use of short-answer questions should do is check that the learning outcomes that he wishes to assess are in fact suited to this form of assessment; if they are not, he should employ some other assessment technique (see booklet on 'Student assessment' for the various alternatives).

Assuming that a short-answer question approach is felt to be appropriate, the next thing that must be decided is what particular type (or types) of item should be used. Here, the best approach is

probably to draw up a detailed list of the various topics and skills that are to be tested and then look at these individually, selecting the type of item that you feel would be most appropriate in each case. Indeed, it is generally a good idea to use the widest possible range of short-answer question types in any given test, since this enables different sorts of skills to be tested and also allows for considerable flexibility in respect of the allocation of marks to the different components of the test. Nor should the choice necessarily be restricted to questions of the short-answer type, since there may well be situations where it would be more appropriate to use questions of the multiple-choice type. Thus, when designing a test to assess skills in the lower-to-middle part of the cognitive domain, the various types of question that are described in this booklet and in the booklet on "Objective questions" should be thought of as forming a single range from which suitable question types can be selected, as appropriate.

Needless to say, full use should be made of any suitable 'ready-made' items (e.g. items held in a central *item bank*) when designing a short-answer question test. Conversely, once you have written a good short-answer question and had it properly validated, it is probably well worth lodging it in a bank of some sort so that it is readily available for future use.

How to write short-answer questions

One of the advantages of basing a test on short-answer questions rather than multiple-choice questions is that the former are generally much easier to write. Nevertheless, such questions do require to be written with some care if they are to be effective assessment instruments, so the following guidelines should be observed.

As in the case of writing a multiple-choice question, the task of writing a short-answer question can be broken down into stages – this time three in number. Let us now look at these in turn.

Stage 1: determining the overall purpose and content of the item
This is a continuation of the process that was described in the previous section, and involves:

- identifying the topic that is to be covered by the item (generally a specific topic from a syllabus of some sort);
- identifying the specific learning outcomes or skills that are to be assessed by the item (generally either *knowledge*, *comprehension*, *application* or *analysis*, or some combination of these.

Stage 2: writing the item

The way in which you set about writing a short-answer question, and the criteria that you try to satisfy while doing so, will obviously depend to a considerable extent on the *style* of the question being written (completion item, unique-answer question or 'open'-question), on the nature of the material being covered by the item, and on the type of learning outcomes or skills being assessed. Nevertheless, there is one basic principle that should be borne in mind at all times: *write the question in such a way that the person answering it is in no doubt as to what is being asked and (provided that he or she is familiar with the material) is in no doubt as to what answer to give*. Obviously, this takes a great deal of skill, since it is all too easy to write such questions in a loose or ambiguous way. This can cause confusion in the mind of the reader, thus making it unlikely that the required answer will always be supplied by the good students.

Bearing the above over-riding principle in mind, there are a number of other basic rules that should be observed when writing short-answer questions.

- Express the item in such a way that only a single, brief answer (or, in the case of a multi-answer item, a series of single, brief answers) is possible. This can be a far from simple task, since what may appear to the writer to be a straightforward, unambiguous question can often be answered in many different ways, depending on how it is interpreted by the reader. The problem is even greater in the case of 'open' short-answer questions that require the respondee to express the answer in his or her own words (like the second example of this type of question given on page 5). Here, it is often a good idea to give some indication of the length of answer required, either by using a formula such as 'state, in roughly 15-30 words, -----' or by providing a space for the response that effectively limits the amount of material that can be written. A further indication of the length of answer expected can often be given via the mark that is awarded to the item or section.
- Whatever the style of item that is to be used, start by expressing it in the form of a direct question. This will increase the likelihood that the problem is clearly stated and that only one answer will be appropriate.
- Express the item in clear, simple language, making it as concise as possible consistent with avoiding looseness or ambiguity. Remember that the purpose of a good short-answer question should be to assess knowledge or skills relating to the subject

matter covered, *not* to assess the testee's ability to decipher the question.

- Express the item in positive form wherever possible, since it has been found that positively-phrased test items tend to measure more important learning outcomes than negatively-phrased items. (Knowing such things as the *best* method or the *most relevant* argument generally has greater educational significance than knowing the *poorest* method or the *least-relevant* argument.)
- If you *have* to use negative wording in an item, use some form of emphasis (e.g. italics, bold type, capitals or underlining) to bring this to the attention of the testee.
- Try to avoid unwittingly providing clues to the required answer – by, for example, incorporating formal prompts in the test (e.g. by using 'a' or 'an' immediately before the missing word in a completion item) or providing answer spaces that are too obviously proportional to the lengths of the required responses.
- Where a numerical answer has to be supplied, indicate the degree of precision expected and (if appropriate) the units in which you want it to be expressed. If you do not do this, you may well have considerable problems in deciding whether the answers supplied are acceptable or not. (Needless to say, this rule should only be applied in cases where knowledge of the degree of precision and/or of the units to be employed is not an integral part of the material being assessed.)

Stage 3: finalising the layout of the item

Once you have decided on the form of words to be used in an item, and have produced any ancillary material required (e.g. an associated figure), some care should be taken in planning the final layout of the item. Again a number of basic rules should be observed.

- Make sure that the item is presented in such a way that the testee is in no doubt as to what he is expected to do and how he is required to indicate his response. If appropriate, provide boxes of suitable size for these responses, or (in the case of 'open' short-answer questions that require more than one or two words) leave a suitable space after each question.

- Make sure that the testee is aware of the total value of the item in terms of marks that can be gained, and (in the case of complex or multi-answer items) show how these marks are divided among the various parts if you feel that this would be of help to the testee.

How to evaluate short-answer questions

Once you have written what you think is a good short-answer question, it is always a good policy to have it evaluated in some way *before* you use it for the purpose for which it has been designed – especially if it is to be used in a formal test or examination of some sort rather than in a more informal setting such as a diagnostic test. The way in which this evaluation is carried out will, of course, depend on the circumstances, and can be anything from simply asking a colleague to read through the item to see if it 'looks OK' to organising a full-scale pre-trial and carrying out a rigorous statistical analysis of the results.

Having an item evaluated by a colleague or validation panel

The most common method of evaluating a short-answer question (or, more usually, a test composed of such questions) is to have it checked by a colleague or validation panel. In order to enable such an evaluation to be carried out in a meaningful and systematic way, it is advisable to present each item on a standard pro-forma of some sort – a pro-forma that not only gives the full text of the item and the model solution, including (if appropriate) the marking scheme, but also states:

- the course, module, syllabus, etc. that the item relates to;
- the purpose for which the item is to be used;
- the particular topic that the item relates to;
- the particular educational skills or outcomes that the item is designed to assess (e.g. knowledge, comprehension, application, etc.);
- the intended level of difficulty, expressed either qualitatively (e.g. 'extremely easy', 'moderately easy', etc.) or quantitatively (e.g. 'expected success rate 70%).'

It should also state the name of the originator of the item and the date of production.

Ideally, the colleague or validation panel carrying out the evaluation should also be provided with an appropriate pro-forma on which to note their conclusions. This should ask for the following information about the item:

- Is the item relevant to the course/module/syllabus to which it relates?
- Is the item style appropriate to the topic being covered and the specific educational skills or outcomes being assessed?
- Does the item present the testees with a clearly-defined task?

- Is the item logically and structurally sound?
- Is the item stated in simple, clear language?
- Is the item free from extraneous clues?
- Is the stated difficulty of the item likely to prove accurate?
- Is this stated difficulty appropriate?

Obviously, any weaknesses or deficiencies identified by the evaluation process should be remedied before the item is put to use – after being subjected to further evaluation, if necessary.

Carrying out a quantitative evaluation of a short-answer item

This can either be done *before* the item is used (by organising a pre-trial with a sample of the population for whom it is designed) or *after* the item has been used (by analysing the responses to the item in the test or examination in which it was incorporated). Whichever method is used, two pieces of information about the item should be obtained.

(i) The *facility value* (or *difficulty index*), which measures the degree of difficulty of the item. As in the case of a multiple-choice item, this is simply the fraction of testees who give the correct answer, and, in the case of a 'good' short-answer question, should normally be between 0.35 and 0.85, with the exact value depending on the purpose for which the item is designed and the intended degree of difficulty.

(ii) The *discrimination index*, which measures the degree to which the item separates the better students from the poorer students. This can be determined in a number of ways, e.g. by calculating the difference between the facility values for the top third of the population (for the test as a whole) and for the bottom third (again for the test as a whole) for the item under consideration. As in the case of a multiple-choice item, the discrimination index for a short-answer question should always be greater than +0.2, and, for a 'good' item, should be at least +0.3.

Clearly, any weaknesses in an item that are brought to light by such an analysis (e.g. an inappropriate facility value or discrimination index) should be remedied before the item is used again. Indeed, if the analysis reveals that the item is a really bad one (e.g. has a negative discrimination index), it should either be completely re-written or discarded.

How to mark short-answer questions

Because they require the testee to *supply* the answer rather than simply select an answer from a number of options provided, short-answer questions are more difficult to mark reliably than multiple-choice questions. However well such a question is formulated, there will always be *some* variation in the answers that are given by the various respondees, thus requiring a certain amount of subjective judgment on the part of the marker. In the case of simple completion items or unique-answer questions, this subjective element can, by careful design, generally be kept fairly low, so that the questions are (for all practical purposes) more or less equivalent to objective questions. In the case of more complex or 'open' short-answer questions, on the other hand, the subjective element can be considerable, and can give rise to *reliability* problems in some cases.

When it comes to assessing the significance of the overall score gained by a testee in a short-answer paper, on the other hand, the situation is much more straightforward than with a multiple-choice paper, since the thorny problem of whether to correct for guessing simply does not arise in this case. Thus, the raw scores gained by the testees can legitimately be taken as true scores, unless, of course, it is felt that some form of standardization would be appropriate.

Further Reading

Readers who require further information on how to write, evaluate and use short-answer questions of the 'quasi-objective' type should find the following two general texts useful.

1. *Objective Testing in Education and Training*, by W. Bonney Rust; Pitman Education Library, London; 1973.
2. *Constructing Achievement tests*, by N. E. Gronlund; Prentice Hall Inc., Englewood Cliffs, New Jersey; 1968.

For teachers of chemistry, physics and mathematics, the following three subject-based texts should also prove useful:

- *The Principles of Objective Testing in Physics*, by J. G. Houston.
- *The Principles of Objective Testing in Chemistry*, by C.U.T. Campbell and W. J. Milne.
- *The Principles of Objective Testing in Mathematics*, by W. G. Fraser and J. M. Gillam.

All three books are published by Heinemann Educational Books, London and Edinburgh.